FibreFix Tutorial 1:
Oriented Crystalline Fibre (E-DNA)

Introduction

A new E form of synthetic DNA was formed from poly(I-I-T).poly(A-C-C) (Leslie *et, al.,* 1980, J Mol Biol. 143, 49-72.). E-DNA gives a well-oriented crystalline fibre diffraction pattern, in this case on film (Figure 1). Its horizontal layer-lines are well sampled by Bragg peaks. Its structure has been analysed with the facilities of the CCP13 software suite for X-ray fibre diffraction data stripping and modelling, particularly the FibreFix program.

Step 1: Scanning the images

The X-ray diffraction pattern of E-DNA has been recorded using a double film pack. The two films, top and bottom, top with strong reflections and bottom with weak intensities, are referred to here as strong and weak. Data from the two films can be merged to increase the dynamic range covered. They were both scanned using a leaf-scan-45 film scanner in the Biological Structure and Function Section, Imperial College London, at 10 μm (2540 dots per inch). Figure 1 shows the strong E-DNA pattern and Figure 2 shows the 'weak' image.

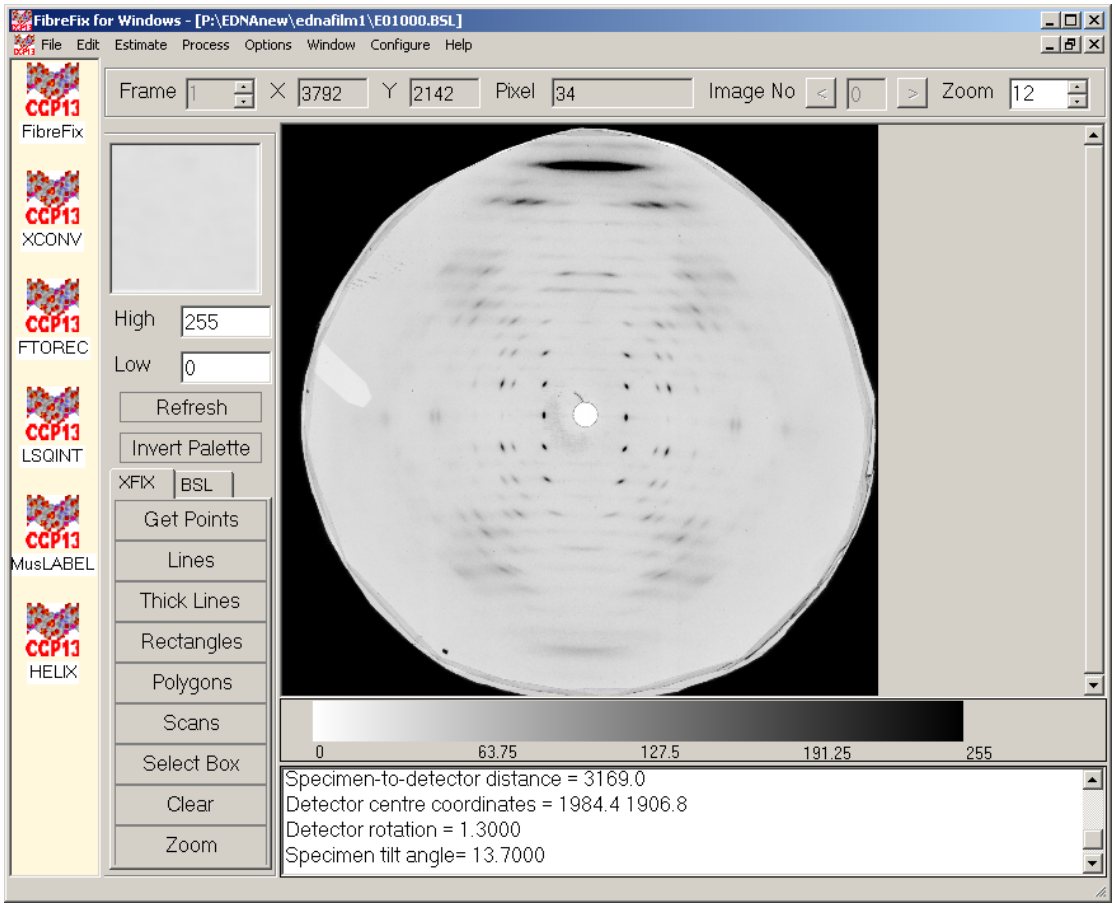

Figure 1: E-DNA strong pattern

Step 2: Image conversion

It is important to convert X-ray images into BSL format for processing in the CCP13 suite. In the past the scanned tiff images would have been loaded into the XCONV program. XCONV will resize the output by linear interpolation of the image pixels and convert it to BSL format. This process leads to an image resolution which is lower than the originally scanned image, but very high quality images are still obtained. The new Windows program FibreFix enables one

to process tiff images without prior conversion into BSL format and it can handle any size of image. It also supports 'Drag and Drop' functionality. Output files from FibreFix are in BSL format.

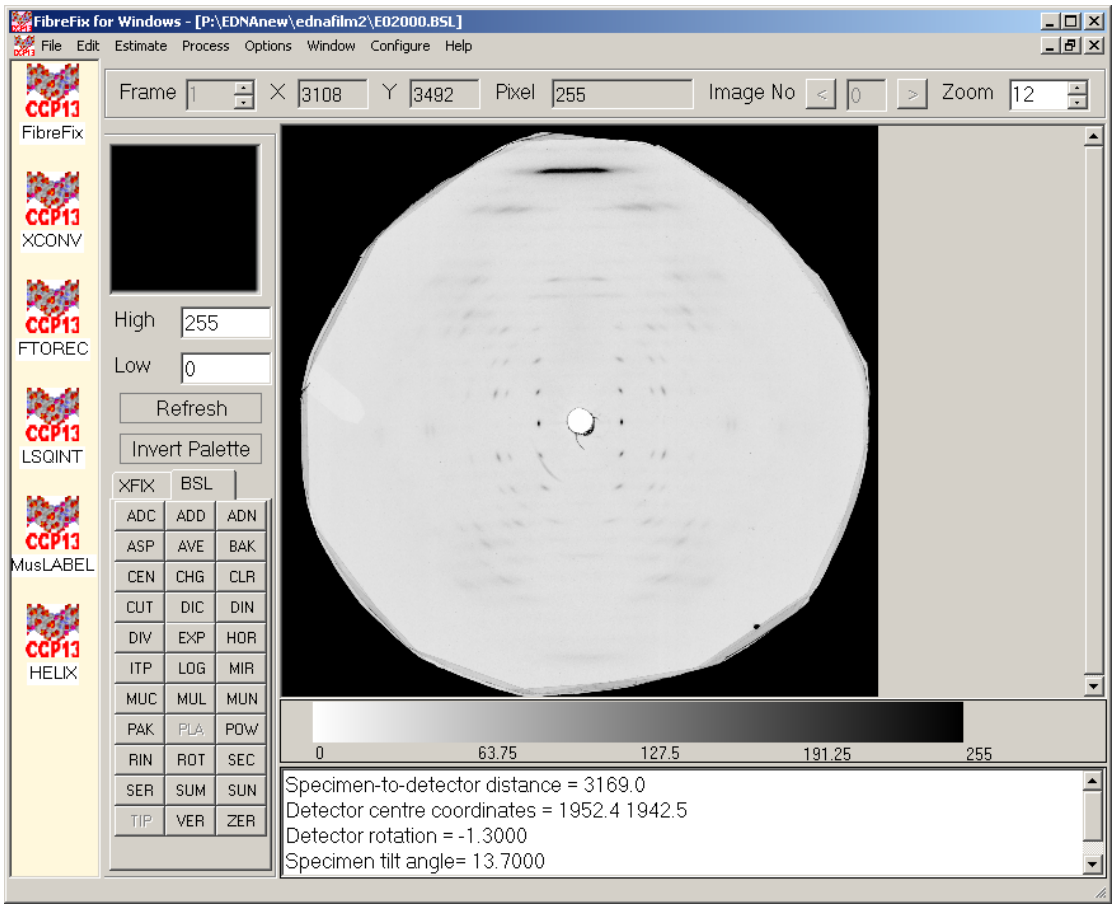

Figure 2: E-DNA weak pattern

Step 3: Parameter calculation

The BSL format images were taken into the FibreFix program for calculating various image parameters. Several parameters need to be determined: these include the pattern centre, the pattern rotation, the camera length and the specimen tilt. First, by clicking on Get Points under XFIX and selecting four equivalent reflections that lie in the four quadrants of the pattern, the centre and rotation can be estimated. This is achieved by carefully setting the cursor on each of the peaks in turn (aided by the enlargement in the top left window of FibreFix) and left-clicking on each using the mouse. Then, by right clicking (Figure 3) and selecting the parameters to determine (centre and rotation) these will be estimated and their values stored in the Parameters Table (Figure 4) found under the drop-down Edit menu. The program estimates the best centre and rotation from these four points and also draws a circle on the points in the image. The specimen to detector distance (camera length) can then be determined from a calibrant peak by inputting the known d spacing for the peak into the Parameter table, found under the Edit drop-down menu, then clicking on Clear and Get Points under XFIX, and clicking on the peak (or peaks) of known spacing. A right click will then open a window saying Estimate Distance. Clicking on this will show the estimated camera length in pixels and put this value into the Parameters table. If the pattern contains several reflections of the same spacing or a ring from a calibration salt, then several points should be selected before right clicking and clicking on estimate distance. If there is no known calibration peak, then the camera length needs to be input in pixels based on the accurate physical measurement of the camera length and the known pixel size in the image (or detector).

All the determined parameters will be stored in the Parameters table. Finally by Clearing once again and once again selecting four equivalent peaks in the four quadrants of the pattern, the specimen tilt angle can be estimated. Then, after choosing an appropriate image area free from artefacts (the backstop etc), the centre and rotation can first be refined, then the tilt and finally all three parameters refined together to get the best values.

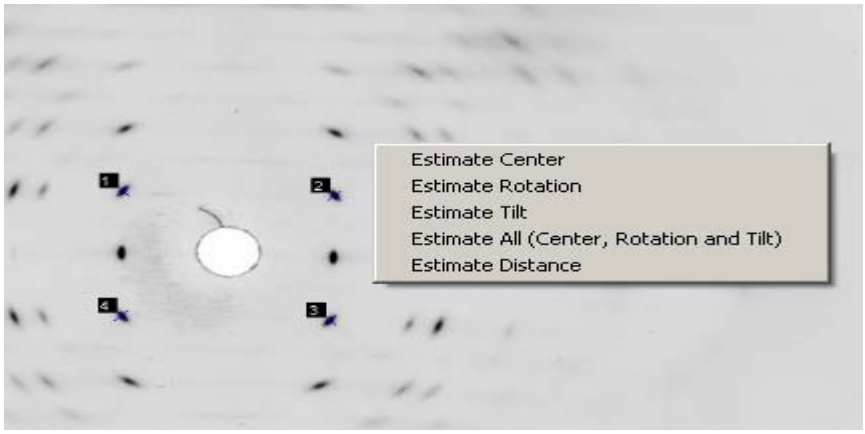

Figure 3: E-DNA parameter estimation

Step 4: Transform into reciprocal space

With knowledge of the rotation, centre, specimen tilt, wavelength and the specimen to film distance, all stored in the Parameters table, the images can be transformed from film space to cartesian reciprocal space using the FTOREC process. To get to FTOREC, go the to Process drop-down menu, and click on the FTOREC option. Figure 4 shows the FTOREC parameter window with some parameters brought forward from the earlier processing and some additional values (e.g. Dmin and Dmin defining the inner and outer radial values in reciprocal space to be used). Figure 5 shows the output of the E-DNA strong film in reciprocal space. A useful feature in FTOREC is the option to output an image corresponding to the standard deviations of the values in the reciprocal space bins. It gives a good check on the accuracy of the parameters obtained in XFIX and several cycles may be needed to optimise these.

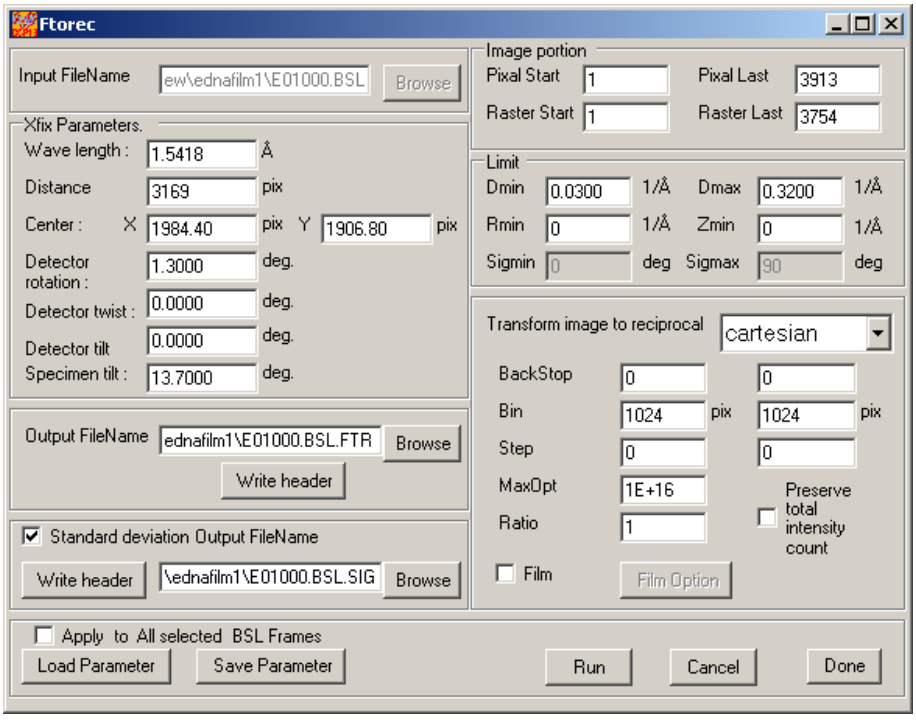

Figure 4: FTOREC window with the parameters for E-DNA

Figure 5: E-DNA strong reciprocal space output image

Step 5: Background subtraction

The Background can be determined either before or after going into FTOREC by using the Background option under the Process drop-down menu. In this case, after transformation of the image to reciprocal space, background subtraction can be carried out using FibreFix with both Paul Langan's "roving window" method and a "smoothed" background through iterative low-pass filtering with different parameter settings. The kind of background estimation to be used depends very much on the particular image being processed and needs to be determined by trial and error. With the E-DNA pattern, by comparing the background and background-subtracted images, the “smoothed background “ process was preferred with thirty cycles for background determination. Figure 6 shows the estimated background and Figure 7 shows the background-subtracted image of the strong pattern.

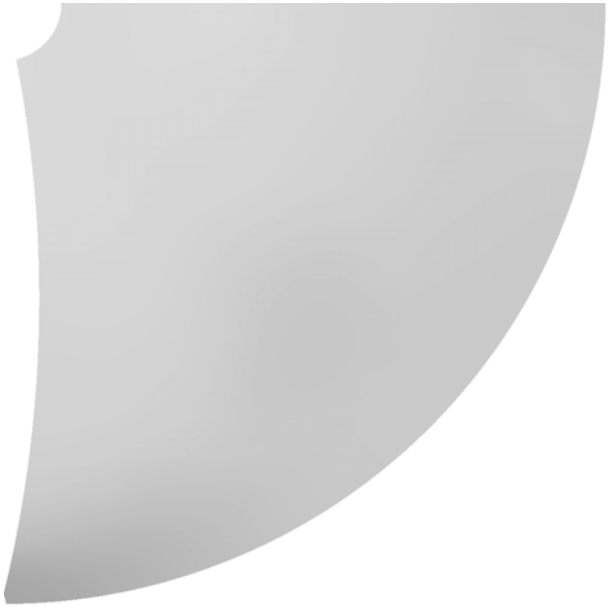

Figure 6: E-DNA strong background image.

Figure 7: E-DNA strong background subtracted image.

Step 6: Fitting the integrated intensity

The integrated intensities of the patterns, after background subtraction, were estimated using the LSQINT process. This was also used to refine the unit cell parameters and the peak profiles of the fitting. LSQINT can selected from the Options drop-down menu only if the image to be processed has been processed through FTOREC. Note that after each process, the images are automatically saved with additional filename extensions. FTOREC images end in .FTR. Since the pattern is Bragg-sampled, fitting was done with a least-squares method. Initial profile parameters were estimated with the NoFit option under the Filenames and Options tab in LSQINT. Once reasonably good profile parameters had been obtained, they were refined to get the best values. Figure 8 shows the refined Parameters used. Fitting and peak profile refinement were repeated until a good R-factor was obtained (it is below 30% here) together with a satisfactory visible output fitted image.

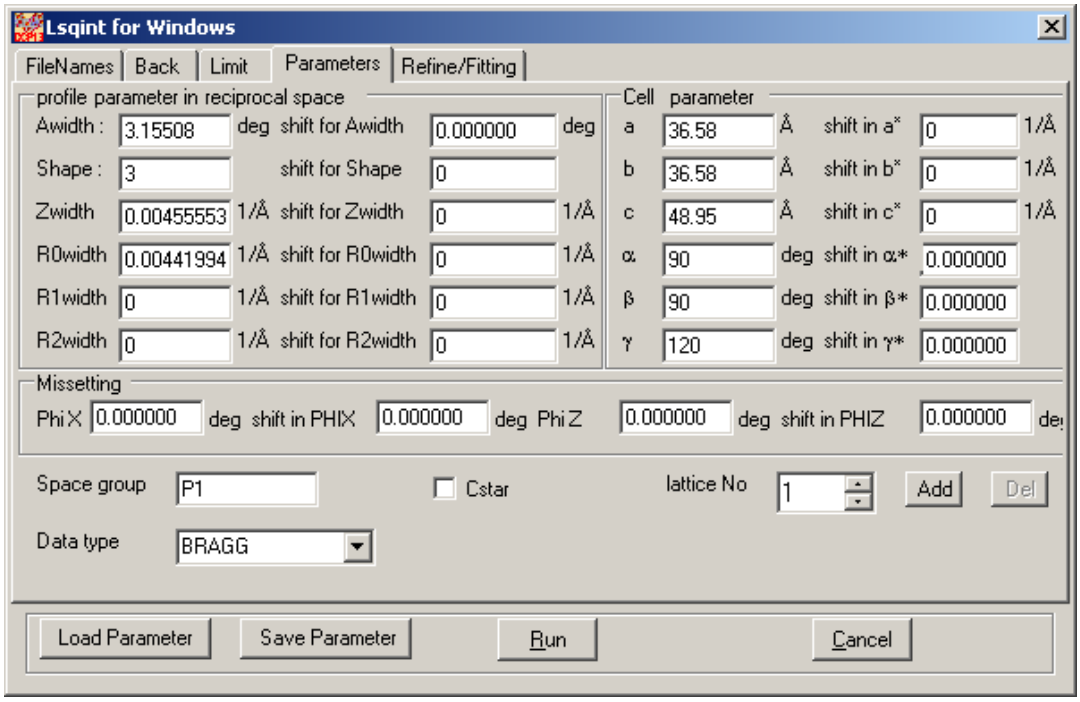

Figure 8: Parameter window with the final refined value.

There are two outputs from LSQINT. The first one is the Intensity output file, formatted with seven columns:

h, k, l, R, M, I, σ

where *h, k, l* correspond to the Miller indices, R is the reciprocal space radius of the sample point, M is the multiplicity, I is the intensity and σ is its standard deviation. Figure 9 shows the first part of the E-DNA Intensity output file.

```
BRAGG
 CELL 36.58 36.58 48.95 90.00 90.00 120.00
 DELTA   3.1250e-04
   1   0   0  3.1570e-02   0  0.0000e+00  0.0000e+00
  -1   1   0  3.1570e-02   0  0.0000e+00  0.0000e+00
   0   1   0  3.1570e-02   6  6.5356e-05  8.5324e-06
  -2   1   0  5.4681e-02   0  0.0000e+00  0.0000e+00
  -1   2   0  5.4681e-02   0  0.0000e+00  0.0000e+00
   1   1   0  5.4681e-02   6  5.6393e-03  1.0653e-04
   2   0   0  6.3140e-02   0  0.0000e+00  0.0000e+00
  -2   2   0  6.3140e-02   0  0.0000e+00  0.0000e+00
   0   2   0  6.3140e-02   6  5.0169e-04  2.6257e-05
```

Figure 9: LSQINT Intensity output.

The second one is a BSL image output file with two frames. The first frame is the observed pattern and the second frame is the fitted intensity. This provides a visible comparison of the outcome of the process and thereby gives an indication of the goodness of fit. Figure10 shows the fitted output image, frame 2 of the output.

Figure10: E-DNA strong LSQINT fitted output image, frame 2 of the output.

Step 7: Structure Determination.

Once good intensity values have been determined from LSQINT the problem is how to determine the associated structure. CCP13 currently uses two alternative programs to do this, LALS and FX-PLOR, each of which also has the facility to minimise the energy of the structure

being studied, along with assessing goodness of fit to the X-ray diffraction data using a crystallographic R-factor. These modelling programs will be implemented in the CCP13 suite and Tutorials on their application will be available on the CCP13 website.

Once the best atomic coordinates have been determined, the structures can be displayed using conventional molecular graphics programs.

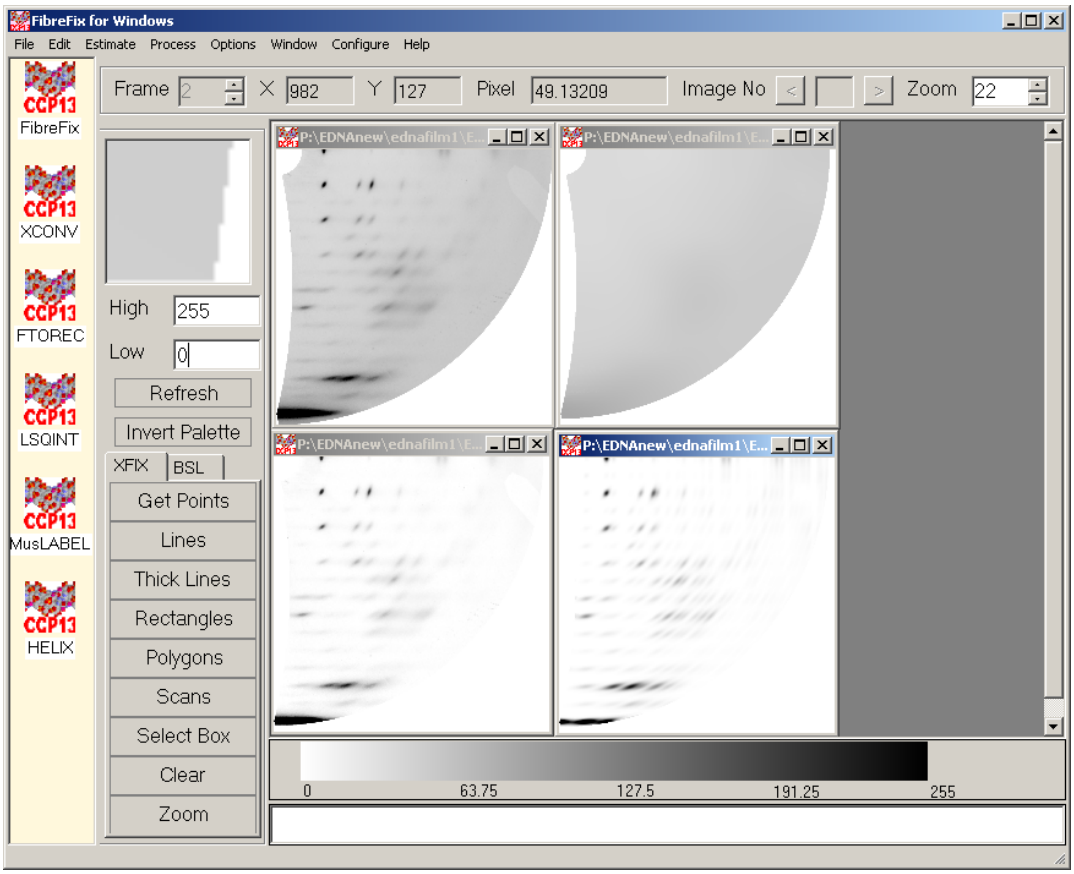

Figure 11: E-DNA FTOREC, background, background subtracted and fitted images.

Tutorial written by:

Ganeshalingam Rajkumar & John Squire

May 19th, 2005